



# Draft Genome Sequence of “*Nocardia suis*massiliense” Strain S-137 (CSUR P4007)

Mustapha Fellag,<sup>a</sup> Anthony Levasseur,<sup>a,b</sup> Jérémy Delerce,<sup>a</sup> Fadi Bittar,<sup>a</sup> Jean-Lou Marié,<sup>b</sup> Bernard Davoust,<sup>a,b</sup> Michel Drancourt<sup>a</sup>

<sup>a</sup>Aix-Marseille Université, MEPHI, IRD, IHU-Méditerranée Infection, Marseille, France

<sup>b</sup>Animal Epidemiology Working Group of the Military Health Service, DRSSA, Toulon, France

**ABSTRACT** “*Nocardia suis*massiliense” strain S-137 isolated from *Sus scrofa* feces exhibits a 9.4-Mb (67.1% GC content) draft genome sequence containing 8,658 protein-coding genes, 66 tRNAs, and 9 rRNAs. *In silico* DNA-DNA hybridization confirmed strain S-137 as representative of a new species, “*Nocardia suis*massiliense,” closely related to *N. tenerifensis* and *N. brasiliensis*.

The genus *Nocardia*, named in honor of the French veterinarian Edmond Nocard (1), contains 115 species in the List of Prokaryotic Names with Standing in Nomenclature (<http://www.bacterio.net/nocardia.html>). Belonging to a suborder of aerobic actinomycetes, the genus *Nocardia* comprises opportunistic pathogens that cause localized and disseminated infections in humans and in animals (2). However, there is little information about diversity and pathogenicity of *Nocardia* species in animals, and only one report has been issued for wild boars (*Sus scrofa*) (3). “*Nocardia suis*massiliense” strain S-137 (CSUR P4007) was isolated from the feces of a wild boar in the military camp at Canjuers, southeastern France (43°38′49″N; 06°27′56″E), using chlorhexidine decontamination and culture on MOD9 medium (4). We analyzed its whole-genome sequence in order to describe its genomic content and help develop molecular identification tools.

*Nocardia* sp. strain S-137 was subcultured on Colombia agar incorporating 5% sheep’s blood (bioMérieux, Marcy l’Etoile, France) at 37°C. DNA extracted using an EZ1 biorobot and an EZ1 DNA tissue kit (Qiagen, Courtaboeuf, France) for <50 µL volume, was quantified by a Qubit assay with a high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 23.1 ng/µL. DNA was then sequenced using MiSeq technology (Illumina, Inc., San Diego, CA, USA) with paired-end and mate pair applications. The index representation for strain S-137 was 6.98%. A total of 866,229 paired-end reads were filtered per read qualities and assembled using SPAdes software (5). Contigs were combined by using SSPACE (6), GapFiller (7), and manual finishing using similarity searches and synteny block detection.

The strain S-137 draft genome was assembled into 25 scaffolds composed of 142 contigs, with a total size of 9,465,473 bp and a GC content of 67.1%. Automatic annotation was performed with Prokka version 1.12 (8). The genome presents two repeat regions and is predicted to encode 8,658 genes, including 8,582 protein-coding genes and 76 RNA genes, including 66 tRNAs, 1 transfer-messenger RNA, 3 5S rRNAs, 3 23S rRNAs, and 3 16S rRNAs. The three 16S rRNA gene copies exhibit a total of 17 nucleotide differences and were compared with the reference *N. brasiliensis* genome (GenBank accession number CP022088), resulting in a similarity score of 98% for one copy, while the two other copies match with 99% similarity.

Based on the 16S rRNA gene sequence proximity, genomes were selected and incorporated into *in silico* DNA-DNA hybridization (DDH) (9). The DDH values were calculated using the Genome-to-Genome Distance Calculator (GGDC) version 2.0 online

**Received** 28 February 2018 **Accepted** 16 March 2018 **Published** 19 April 2018

**Citation** Fellag M, Levasseur A, Delerce J, Bittar F, Marié J-L, Davoust B, Drancourt M. 2018. Draft genome sequence of “*Nocardia suis*massiliense” strain S-137 (CSUR P4007). *Genome Announc* 6:e00212-18. <https://doi.org/10.1128/genomeA.00212-18>.

**Copyright** © 2018 Fellag et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Michel Drancourt, [michel.drancourt@univ-amu.fr](mailto:michel.drancourt@univ-amu.fr).

tool (10). This analysis yielded 31.2% sequence similarity with the type strain *N. tenerifensis* NBRC 101015 (NCBI reference sequence NZ\_BAGH000000000), 27.5% sequence similarity with *N. brasiliensis* ATCC 700358 (NC\_018681), 24.9% sequence similarity with *N. asiatica* NBRC 100129 (NZ\_BAFS000000000), 24.8% sequence similarity with *N. abscessus* NBRC 100374 (NZ\_BAFP000000000), and 24.5% sequence similarity with *N. arthritidis* NBRC 100137 (NZ\_BDBB000000000). These data indicate that strain S-137 is representative of a new species, "*Nocardia suismassiliense*," that is closely related to *N. tenerifensis* and *N. brasiliensis*.

These data illustrate the importance of whole-genome sequencing in the taxonomic study and identification of new *Nocardia* species, since 16S rRNA gene sequencing may be insufficient for distinguishing closely related species.

**Accession number(s).** The genome sequence reported here has been deposited at EMBL/GenBank under the accession number [OIFR000000000](#). The version described here is the first version, OIFR01000000.

## ACKNOWLEDGMENTS

This work was supported by the French Government under the "Investissements d'avenir" (Investments for the Future) program managed by the Agence Nationale de la Recherche (Méditerranée Infection 10-IAHU-03). M.F. benefited from a PhD grant from the Fondation Méditerranée Infection.

## REFERENCES

1. Zhi X-Y, Li W-J, Stackebrandt E. 2009. An update of the structure and 16S rRNA gene sequence-based definition of higher ranks of the class *Actinobacteria*, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa. *Int J Syst Evol Microbiol* 59:589–608. <https://doi.org/10.1099/ijs.0.65780-0>.
2. McNeil MM, Brown JM. 1994. The medically important aerobic actinomycetes: epidemiology and microbiology. *Clin Microbiol Rev* 7:357–417. <https://doi.org/10.1128/CMR.7.3.357>.
3. Matos AC, Dias AP, Morais M, Matos M, Pinto ML, Coelho AC, Figueira L. 2015. Granulomatous lymphadenitis caused by *Nocardia* species in hunted wild boar (*Sus scrofa*) in Portugal. *Vet Rec* 177:103–104. <https://doi.org/10.1136/vr.h3988>.
4. Asmar S, Chatellier S, Mirande C, van Belkum A, Canard I, Raoult D, Drancourt M. 2015. A novel solid medium for culturing *Mycobacterium tuberculosis* isolates from clinical specimens. *J Clin Microbiol* 53: 2566–2569. <https://doi.org/10.1128/JCM.01149-15>.
5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
6. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
7. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with Gap-Filler. *Genome Biol* 13:R56. <https://doi.org/10.1186/gb-2012-13-6-r56>.
8. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
9. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 106: 19126–19131. <https://doi.org/10.1073/pnas.0906412106>.
10. Auch AF, Jan M, Klenk H-P, Göker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2:117. <https://doi.org/10.4056/sigs.531120>.